บทที่ 3
ระเบียบวิธีวิจัย

3.1 กรอบแนวคิดในการศึกษา และแบบจำลอง

3.1.1 กรอบแนวคิดในการศึกษา

ภาพที่ 3.1 กรอบแนวคิดในการศึกษา

3.1.2 แบบจำลองในการศึกษา

แบบจำลองทางคณิตศาสตร์ มีองค์ประกอบ 3 ส่วน ได้แก่ 1) ตัวแปร (Variables) 2) สมการที่แสดงความสัมพันธ์ของตัวแปร 3) ข้อสมมติ (Assumptions) โดยการศึกษาครั้งนี้ใช้แบบจำลอง Logit Model มีรายละเอียด ดังนี้

Logit Model หรือการวิเคราะห์ถดถอยโลจิท (Logit Regression Analysis) คือเทคนิคการวิเคราะห์ตัวแปรหลายตัว เป็นเทคนิคทางสถิติที่สามารถควบคุมตัวแปรได้หลายตัว พร้อม ๆ กัน เพราะในการวิจัยจำเป็นต้องควบคุมตัวแปรอีกหลายตัวที่มีอิทธิพลต่อความสัมพันธ์ระหว่างตัวแปรที่เป็นเป้าหมายของการวิจัย ดังนั้นการควบคุมปัจจัยที่มีผลต่อตัวแปรตามได้มากเท่าใด ข้อสรุปที่ได้เกี่ยวกับความสัมพันธ์ระหว่างตัวแปรต้น และตัวแปรตามจะถูกต้องมากขึ้นเท่านั้น เพราะได้ขจัดข้อโต้แย้ง 3 ประการ ได้แก่ 1) ความสัมพันธ์ระหว่างตัวแปร 2 ตัว ไม่ใช่ความสัมพันธ์ลวง (spurious relationship) เพราะได้มีการนำตัวแปรอื่นมาร่วมพิจารณาแล้ว 2) ตามหลักเหตุผล และทฤษฎี ตัวแปร 2 ตัวมีความสัมพันธ์กัน แต่หากการวิจัยพบว่าไม่มีความสัมพันธ์กัน อาจเกิดจากการไม่ได้ควบคุมตัวแปรอื่น ๆ หรือนำตัวแปรอื่นมาวิเคราะห์ร่วมด้วย 3) ทิศทางความสัมพันธ์ หากไม่มีการพิจารณาตัวแปรหลายตัวร่วมกัน อาจทำให้ทิศทางความสัมพันธ์ของตัวแปรตาม และตัวแปรต้นที่เป็นเป้าหมายของการวิจัยมีทิศทางตรงกันข้ามกับหลักเหตุผล และทฤษฎีได้

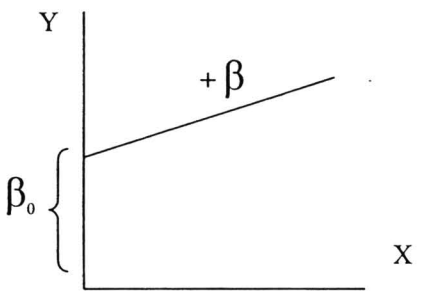
เทคนิคการวิเคราะห์ตัวแปรหลายตัว นอกจากการวิเคราะห์ถดถอยโลจิท หรือการวิเคราะห์สหสัมพันธ์เชิงซ้อน (Multiple Correlation Analysis) ได้แก่ การวิเคราะห์ถดถอยพหุแบบปกติ มีวิธีการประมาณค่าพารามิเตอร์ คือ OLS (Ordinary Least Squares Method, OLS) แต่การวิเคราะห์ถดถอยพหุแบบปกติ แตกต่างจากการวิเคราะห์ถดถอยโลจิท คือ การวิเคราะห์ถดถอยโลจิท ตัวแปรตามมีค่าเป็น 1 กับ 0 ส่วนตัวแปรต้นมีค่าเป็น 1 กับ 0 หรือตัวแปรเชิงปริมาณ ส่วนการวิเคราะห์ถดถอยพหุแบบปกติ ทั้งตัวแปรตาม และตัวแปรต้นเป็นตัวแปรเชิงปริมาณ คือ เป็นข้อมูลช่วง หรืออัตราส่วน

การวิเคราะห์ถดถอยโลจิท จะมีการนำเสนอในรูปแบบของสมการ ดังนี้

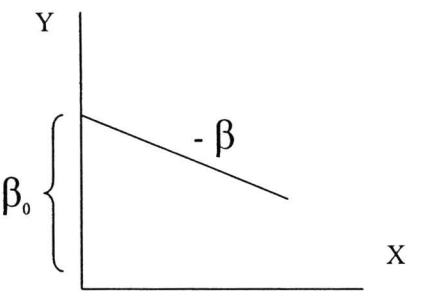
$$\log\left(\frac{P(y=1)}{1-P(y=1)}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots\ldots\ldots\ldots\ldots\ldots$$

โดยที่ตัวแปรตามอยู่ด้านซ้ายมือ และตัวแปรต้นอยู่ด้านขวามือ ตัวแปรต้นจะมีค่าสัมประสิทธิ์ (β) แสดงลักษณะความสัมพันธ์ระหว่างตัวแปรต้น กับตัวแปรตามว่าเป็นไปในทิศทางใด + หรือ – และค่าสัมประสิทธิ์ (โดยไม่คำนึงเครื่องหมาย) ยังแสดงปริมาณความสัมพันธ์ระหว่างตัวแปรต้น กับตัวแปรตามว่ามีค่ามากน้อยเพียงใด ถ้ามีค่ามากก็แสดงว่าตัวแปรต้นตัวนั้นมีความสัมพันธ์กับตัวแปรตามมาก เมื่อพิจารณากราฟจะพบว่าค่าสัมประสิทธิ์ คือค่าความลาดชัน (slope) นั่นคือ ความสัมพันธ์ระหว่างตัวแปรต้น กับตัวแปรตามยิ่งมีค่ามาก ความลาดชันก็จะมาก

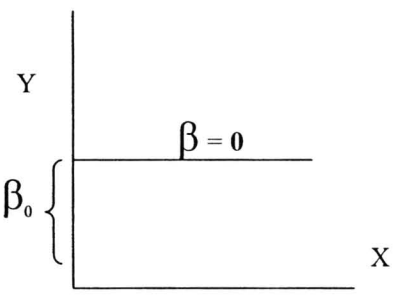
หมายความว่า เมื่อค่าของตัวแปรต้นเปลี่ยนแปลง ค่าของตัวแปรตามก็จะเปลี่ยนแปลงมาก ส่ วนค่า β_0 คือ ค่าคงที่ (constant) หรือจุดตัด (intercept) หรือค่าจุดเริ่มต้น (origin) ดังรูป



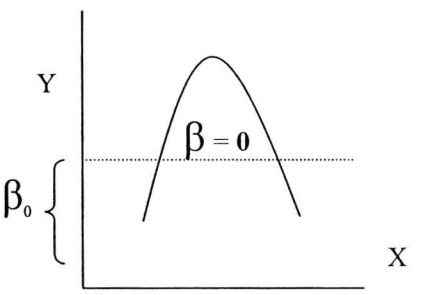
ภาพที่ 3.2 ความสัมพันธ์เชิงบวก



ภาพที่ 3.3 ความสัมพันธ์เชิงลบ



ภาพที่ 3.4 ความสัมพันธ์ = 0



ภาพที่ 3.5 ความสัมพันธ์เส้นโค้ง

ค่าสัมประสิทธิ์เป็นบวก หมายถึง การผันแปรของตัวแปรตาม และตัวแปรต้นเป็นไปในทิศทางเดียวกัน คือ ถ้าค่าของตัวแปรต้นเพิ่มขึ้น ค่าของตัวแปรตามจะเพิ่มขึ้นตามไปด้วย

ค่าสัมประสิทธิ์เป็นลบ หมายถึง การผันแปรของตัวแปรตาม และตัวแปรต้นเป็นไปในทิศทางตรงกันข้าม คือ ถ้าค่าของตัวแปรต้นเพิ่มขึ้น ค่าของตัวแปรตามจะลดลง

ความสัมพันธ์ = 0 หมายถึง ตัวแปรต้น ไม่มีความสัมพันธ์กับตัวแปรตาม ไม่ว่า X จะมีค่าเพิ่มขึ้นเท่าใด Y มีค่าเท่าเดิม

ความสัมพันธ์เชิงเส้นโค้ง หมายถึง X มีความสัมพันธ์เชิงเส้นโค้งกับ Y ทำให้ถ้าใช้เทคนิคการวิเคราะห์ถดถอยมาใช้ X จะไม่มีความสัมพันธ์กับ Y เพราะ ค่าสัมประสิทธิ์ถดถอยมีค่าเป็น 0 ซึ่งในความเป็นจริง X กับ Y มีความสัมพันธ์กัน แต่สัมพันธ์เชิงเส้นโค้ง

ความสัมพันธ์มาก – น้อย ระหว่างตัวแปรต้น และตัวแปรตาม พิจารณาจากค่าตัวเลข ซึ่งส่วนมากเป็นจุดทศนิยม ยิ่งค่าสัมประสิทธิ์มีค่ามากเท่าใด (โดยไม่คำนึงถึงเครื่องหมายบวก หรือลบ) ก็แสดงว่าตัวแปรต้น และตัวแปรตามมีความสัมพันธ์กันมากเท่านั้น

แนวคิดหลักของการถดถอยโลจิสติค คือ โลจิท (logit) หมายถึง ค่าล๊อกธรรมชาติ (natural logarithm) ของอัตราส่วนของโอกาส (odd ratio) โดยการหาค่าโลจิท (logit) ของอัตราส่วนโอกาส และเปลี่ยนให้เป็นค่าล็อกธรรมชาติ (ln) ซึ่งคือค่าสัมประสิทธิ์ถดถอยยกกำลัง (Expβ) ของตัวแปรอิสระ หรือตัวพยากรณ์ (Predictor) สำหรับตัวแปรกลุ่มทวิ (Binary) ค่าอัตราส่วนโอกาส (odd ratio) จะเท่ากับ e (ล๊อกฐานธรรมชาติ) ยกกำลังด้วยค่า β หรือ (Expβ)

ความหมายของ (expβ) หรือ e^{β} สรุปได้ 3 ลักษณะ ดังนี้

1. ถ้า $\beta > 0$ จะทำให้ $e^{\beta} > 1$ หมายความว่า ค่า Odds เพิ่มขึ้น หรือโอกาสที่จะเกิดเหตุการณ์ที่สนใจเพิ่มมากขึ้น

2. ถ้า $\beta < 0$ จะทำให้ $e^{\beta} < 1$ หมายความว่า ค่า Odds ลดลง หรือโอกาสที่จะเกิดเหตุการณ์ที่สนใจลดลง

3. ถ้า $\beta = 0$ จะทำให้ $e^{\beta} = 1$ หมายความว่า ค่า Odds ไม่เพิ่มขึ้น หรือลดลง

Odd Ratio จะแสดงถึงโอกาสที่จะเกิดเหตุการณ์เป็นกี่เท่าของโอกาสที่จะไม่เกิด ถ้าค่า Odd Ratio มากกว่า 1 แสดงว่า โอกาสการเกิดเหตุการณ์มากกว่าการไม่เกิดเหตุการณ์ เช่น Odd Ratio = 3.5 แสดงว่า โอกาสที่จะเกิดเหตุการณ์เป็น 3.5 เท่าของโอกาสที่จะไม่เกิด

e^{β} หรือ Exp(β) เป็นค่าที่แสดงถึงการเปลี่ยนแปลงของ Odd Ratio เมื่อตัวแปรทำนายเปลี่ยนไป 1 หน่วย ถ้าค่า Exp(β) มีค่ามากกว่า 1 แสดงว่า เมื่อตัวแปรทำนายมีค่าเพิ่มขึ้น จะช่วยเพิ่มโอกาสของการเกิดเหตุการณ์ (Y→ 1) สำหรับตัวแปรทำนายที่เป็นทวิ เช่น เพศ จะช่วยเพิ่มโอกาสของการเกิดเหตุการณ์ (Y = 1) สูงขึ้นกี่เท่าของการเกิดเหตุการณ์ (Y = 0)

แต่ถ้าค่า Exp(β) มีค่าน้อยกว่า 1 แสดงว่า เมื่อตัวแปรทำนายมีค่าเพิ่มขึ้น จะช่วยลดโอกาสของการเกิดเหตุการณ์ (Y→ 0) สำหรับตัวแปรทำนายที่เป็นทวิ เช่น เพศ จะช่วยลดโอกาสของการเกิดเหตุการณ์ (Y = 1) สูงขึ้นกี่เท่าของการเกิดเหตุการณ์ (Y = 0)

นอกจากนี้สิ่งที่ต้องพิจารณาหลังจากได้ผลการวิเคราะห์การถดถอยแบบโลจิสติคแล้ว คือการพิจารณาว่าแบบจำลองเข้ากับข้อมูลได้ดี (goodness of fit) หรือไม่มากน้อยเพียงใด มีประเด็นที่ต้องพิจารณา ดังนี้

1) การทดสอบทางสถิติเกี่ยวกับตัวแปรอิสระแต่ละตัว ได้แก่ การพิจารณาค่าสถิติพื้นฐาน เช่น ค่าเฉลี่ย (Mean) ค่ามัธยฐาน (Median) ค่าสูงสุด (Maximum) ค่าต่ำสุด (Minimum) ค่าส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation) และเพื่อทดสอบว่าการวิเคราะห์ด้วยแบบจำลอง Logit Model ในการศึกษาเป็นไปตามสมมติฐานของสมการโลจิท ต้องพิจารณาเงื่อนไข ดังนี้

1.1) ตัวแปรตาม (Y) ต้องเป็น Binary Response ตัวแปรต้น (X) เป็น Dummy Variable / Interval / Ratio Scale ก็ได้

1.2) ค่าเฉลี่ยของตัวคลาดเคลื่อน = 0 [E(€i) = 0] ถ้าตัว various ของตัว error term มีค่าไม่คงที่ แสดงว่ามีปัญหา Heteroscedasticity ซึ่งสามารถทดสอบได้ด้วยวิธี White's test

1.3) ค่าความคลาดเคลื่อนไม่มีความสัมพันธ์กันเอง [Cov(€i€j)=0] พิจารณาจากค่า Durbin Watson ซึ่งใช้กับข้อมูลที่เป็นลักษณะของ Time Series

1.4) ตัวแปรอิสระ กับค่าความคลาดเคลื่อนต้องเป็นอิสระแก่กัน คือ ถ้าค่าความคลาดเคลื่อนทำให้ตัวแปรอิสระมีผลต่อตัวแปรตามน้อย แสดงว่ามีปัญหา ซึ่งสามารถทดสอบได้โดยการพิจารณาค่าการกระจายตัวแบบปกติ (Jarque – Bera: JB) เป็นการทดสอบการแจกแจงปกติของตัวคลาดเคลื่อน (Normality Test) ใช้กับการทดสอบตัวอย่างขนาดใหญ่ที่มีตัวอย่างตั้งแต่ 30 ตัวอย่างขึ้นไป

1.5) ตัวแปรอิสระจะต้องไม่มีความสัมพันธ์กันเอง คือ การทดสอบปัญหา Multicolinearity

1.6) จำนวนตัวอย่างต้องมีอย่างน้อย หรือเท่ากับ 30*P [n≥30*P] [P = Parameter]

2) แบบจำลองในภาพรวม พิจารณาค่า – 2LL (log likelihood) ถ้าแบบจำลองเหมาะสมดีมากกับข้อมูล ค่า – 2LL จะเท่ากับ 0 (ค่า – 2 LL เป็นค่าคำนวณแบบจำลองที่มีค่าคงที่ ซึ่งมีจุดตัดเพียงจุดเดียว ใช้เป็นฐานในการเปรียบเทียบกับแบบจำลองที่มีตัวแปรต้นเพิ่มขึ้น) ค่าของ – 2LL ลดลง แสดงว่าตัวแปรต้นตัวนั้นทำให้แบบจำลองเข้ากับข้อมูลได้มากขึ้น

นอกจากค่า –2LL สามารถพิจารณาแบบจำลองในภาพรวมด้วยสถิติทดสอบ 4 ตัว ได้แก่ (1) อัตราส่วนความเป็นไปได้ (likelihood ratio) (2) คะแนน (score) (3) สถิติทดสอบวอลด์ (Wald tests) และ (4) สถิติทดสอบการเข้ากับข้อมูลได้ดีฮอสเมอร์แอนด์เลมโชว์ (Hosmer & Lemeshow) โดยทั่วไปสถิติทั้ง 4 ตัวจะให้ผลคล้ายคลึงกัน แต่ถ้าให้ผลแตกต่างกันควรใช้อัตราส่วนความเป็นไปได้ (likelihood ratio) และการทดสอบคะแนน (score) เท่านั้น หรือพิจารณาจากค่า McFadden R – Square ซึ่งโดยปกติจะมีค่าไม่เกิน 0.40

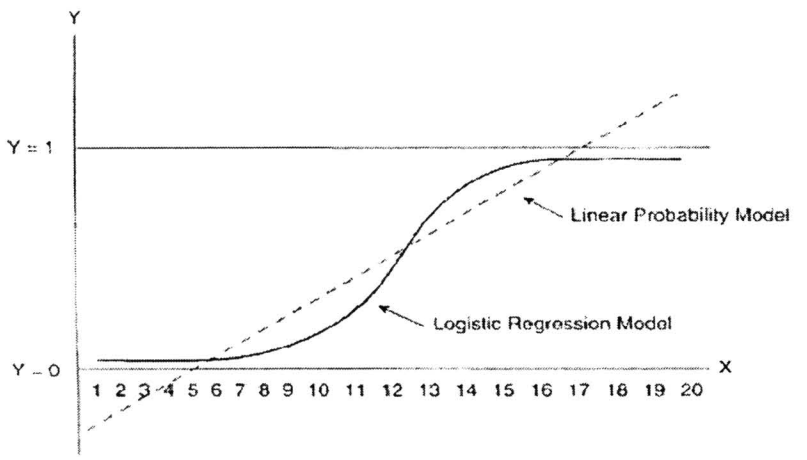
3) ความเหมาะสมใช้ได้ดีของแบบจำลอง (goodness of fit statistics) สำหรับแบบจำลอง Logit Model ให้ความสำคัญค่อนข้างน้อยกับ goodness of fit statistics แต่ให้ความสำคัญมากกับการวิเคราะห์ คือ เครื่องหมายที่แสดงความสัมพันธ์ระหว่างตัวแปรต้น กับตัวแปรตาม (β) เป็น + หรือ – หรือ 0

4) การประเมินความถูกต้องของโอกาส / ความน่าจะเป็นที่ได้จากการพยากรณ์ (predicted probabilities) พิจารณาจาก Classification Table หรือ Overall Correct Classification

แบบจำลอง Logit Model ตัวแปรตามมีค่าเพียงสองค่าคือ 1 กับ 0 เท่านั้น ดังนั้นถ้า

ตัวแปรต้น X มีค่าน้อย ๆ ก็ทำให้ตัวแปรตาม Y มีค่าเท่ากับ 0

ถ้าตัวแปรต้น X มีค่ามาก ๆ ก็ทำให้ตัวแปรตาม Y มีค่าเท่ากับ 1



ภาพที่ 3.6 ฟังก์ชั่นเส้นโค้งของ Logistic Regression

ที่มา: http:// www.tourismlogistic.com

ในทางคณิตศาสตร์สามารถเขียนฟังก์ชั่น Logistics ได้ดังนี้

$$P(Y = 1) = \frac{e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots}}{1 + e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots}}$$

$$= \frac{e^{bx}}{1 + e^{bx}}$$

$$= \frac{1}{1 + e^{-bx}}$$

ฟังก์ชั่นนี้มีความหมายว่า เมื่อใส่ค่า X (ตัวแปรต้น) จะทำให้ค่า Y (ตัวแปรตาม) ที่ได้มีค่าระหว่าง 0 กับ 1 โดยไม่รวมค่า Y = 0 และ Y = 1

การวิเคราะห์ถดถอยโลจิสติค จะมีการแปลงค่าที่ประมาณได้ให้อยู่ช่วง 1 และ 0 ดังนั้น ถ้าตัวแปรตามมีค่าเป็น 1, 2, 3 ซึ่งเป็นลักษณะของตัวแปรแบบ Nominal Scale ทำให้ไม่สามารถนำมาเข้าสมการวิเคราะห์ถดถอยได้ เพราะไม่สามารถบวกลบคูณหารได้เหมือนตัวเลขปกติ ดังนั้นต้องทำให้เป็นตัวแปรหุ่นก่อน โดยที่จำนวนตัวแปรหุ่นของตัวแปรแบบ Nominal Scale แต่ละตัวจะน้อยกว่าจำนวนกลุ่มย่อยของตัวแปรนั้นอยู่ 1 เสมอ

สมการวิเคราะห์ถดถอยจะมีตัวแปรหุ่นกี่ตัวก็ได้ ตราบใดที่ยังมีอัตราความเป็นอิสระมากพอสมควร คือจำนวนตัวแปรตามยังน้อยกว่าจำนวนหน่วยวิเคราะห์ (กลุ่มตัวอย่าง)

ข้อจำกัดของการวิเคราะห์การถดถอยโลจิท ได้แก่

1) ความเป็นเส้นตรง (linearity) ของความสัมพันธ์ระหว่างตัวแปรต้น และตัวแปรตาม ดังแสดงในภาพที่ 3.2, 3.3, 3.4 ดังนั้นหาก X กับ Y มีความสัมพันธ์กันแบบเส้นโค้ง การวิเคราะห์ความถดถอยโลจิท จะทำให้ได้ค่าสัมประสิทธิ์ = 0 ซึ่งแสดงว่า X กับ Y ไม่มีความสัมพันธ์กัน ทั้งที่ในความเป็นจริงแล้วมีความสัมพันธ์กัน ดังนั้นต้องปรับสมการให้เป็นสมการเส้นตรงก่อนจึงใช้การวิเคราะห์ความถดถอยโลจิทได้

2) ความสัมพันธ์กันมากระหว่างตัวแปรอสิระด้วยกัน (Multicollinearity) วัตถุประสงค์การวิเคราะห์การถดถอยโลจิท คือ ศึกษาอิทธิพลตัวแปรต้นแต่ละตัวที่มีต่อตัวแปรตาม และอิทธิพลของตัวแปรต้นทั้งหมดต่อตัวแปรตาม และเมื่อได้แบบแผนความสัมพันธ์ระหว่างตัวแปรแล้วก็นำไปประมาณค่าของโอกาส หรือความน่าจะเป็น (probabilities) ที่หน่วยวิเคราะห์แต่ละหน่วยจะเป็นสมาชิกของกลุ่มใดกลุ่มหนึ่ง ดังนั้น หากตัวแปรต้นแต่ละตัวสัมพันธ์กันก็ไม่สามารถแยกอธิบายอิทธิพลของตัวแปรต้นแต่ละตัวต่อตัวแปรตามได้ ซึ่งหากสมการมีปัญหา Multicollinearity สามารถแก้ปัญหาได้ 3 แบบ คือ

2.1) ตัดตัวแปรใดตัวแปรหนึ่งที่สัมพันธ์กันมากออกไป

2.2) สร้างตัวแปรใหม่จาก 2 ตัวแปรตาม แล้วใช้ตัวแปรใหม่นั้นแทน

2.3) การขจัดผลของตัวแปรตัวหนึ่งออกจากตัวแปรอีกตัวหนึ่ง

3) การพิจารณาค่าเปอร์เซ็นต์การทำนาย เป็นการเปรียบเทียบจำนวนกรณีที่เป็นค่าสังเกต (observed cases) กับจำนวนกรณีที่เป็นค่าทำนาย หรือจำแนกได้โดยใช้แบบจำลองว่าถูกต้องร้อยละเท่าใด ยิ่งถูกต้องมาก แบบจำลองยิ่งใช้ได้ดีมาก ค่าของโอกาสที่กำหนดไว้ล่วงหน้า (prior probabilities) คือ ค่าของโอกาสของแต่ละกรณีที่จะถูกจำแนกได้อย่างถูกต้องในกลุ่มใดกลุ่มหนึ่ง ค่าดังกล่าวนี้ คืออัตราส่วนของขนาดของประชากรของแต่ละกลุ่ม ส่วนค่าของโอกาสที่ได้ภายหลัง (posterior probabilities) คือค่าของโอกาสที่แต่ละกรณีจะถูกกำหนด โดยแบบจำลองที่ใช้ เมื่อนำแบบจำลองไปใช้จะถูกต้องได้ร้อยละเท่าใดของตัวอย่างทั้งหมด

การศึกษาครั้งนี้ใช้แบบจำลอง Logit Model โดยใช้โปรแกรม EViews ซึ่งเป็นโปรแกรมที่นิยมใช้ในการวิเคราะห์ทางเศรษฐมิติ โดยสามารถคำนวณข้อมูลที่เป็นข้อมูลอนุกรมเวลา (Time series data) และข้อมูลประเภทตัดขวาง (Cross section data) ได้ สำหรับการศึกษาปัจจัยทางเศรษฐสังคมที่มีผลต่อการเกิดภาวะโรคอ้วนของนักศึกษาในมหาวิทยาลัย จังหวัดเชียงใหม่เป็นการศึกษาข้อมูลประเภทภาคตัดขวาง (Cross section data) โดยมีสมการในการวิจัย ดังนี้

Y = f {S, W, HI, FI, EF, EM, OF, OM, R, PT, SF, L, CM, CL, CE, KM, KL, KE, F, D, E, OP, I, M}

ตัวแปรตาม

Y = 1 BMI $\geq$ 23 = อ้วน

0 BMI < 23 = ไม่อ้วน

ตัวแปรต้น

S = เพศ

W = การรับรู้น้ำหนักของตัวเอง

HI = ประกันสุขภาพเอกชน

FI = รายได้ต่อเดือนของครอบครัว

EF = ระดับการศึกษาของบิดา

EM = ระดับการศึกษาของมารดา

OF = อาชีพของบิดา

OM = อาชีพของมารดา

R = จำนวนเงินที่ได้รับต่อเดือน

PT = รายได้เสริมจากการทำงานพิเศษ

SF = ค่าใช้จ่ายสำหรับอาหาร และเครื่องดื่ม

L = สถานที่พักอาศัย

CM = ลักษณะของอาหารเช้า

CL = ลักษณะของอาหารกลางวัน

CE = ลักษณะของอาหารเย็น

KM = ชนิดของอาหารเช้า

KL = ชนิดของอาหารกลางวัน

KE = ชนิดของอาหารเย็น

F = การบริโภคอาหารฟาสฟู้ด

DR = การบริโภคเครื่องดื่ม

E = การออกกำลังกาย

OP = ความเต็มใจจ่ายในการรักษาโรคอ้วน

I = ข้อมูลข่าวสาร

M = พฤติกรรมผู้บริโภคแบบ Myopic View

Y เป็นตัวแปรหุ่น

การมีภาวะน้ำหนักเกิน และโรคอ้วน BMI $\geq$ 23 กำหนดให้เป็น 1

การไม่เป็นโรคอ้วน BMI < 23 กำหนดให้เป็น 0

S เป็นตัวแปรหุ่น

เพศชาย กำหนดให้เป็น 1

เพศหญิง กำหนดให้เป็น 0

W เป็นตัวแปรหุ่น (การรับรู้น้ำหนักของตัวเอง ทำการแปรผลน้ำหนักของผู้ตอบแบบสอบถามเปรียบเทียบกับค่า BMI)

การรับรู้น้ำหนักของตัวเองไม่ถูกต้อง กำหนดให้เป็น 1

การรับรู้น้ำหนักของตัวเองถูกต้อง กำหนดให้เป็น 0

HI เป็นตัวแปรหุ่น

มีประกันสุขภาพเอกชน กำหนดให้เป็น 1

ไม่มีประกันสุขภาพเอกชน กำหนดให้เป็น 0

FI เป็นตัวแปรเชิงปริมาณ คือ รายได้ต่อเดือนของครอบครัว (บาท/เดือน)

EF เป็นตัวแปรเชิงกลุ่ม คือ ระดับการศึกษาของบิดา

ประถมศึกษาหรือต่ำกว่า กำหนดให้เป็น 0

มัธยมศึกษาตอนต้น กำหนดให้เป็น 1

มัธยมปลาย/ปวช. กำหนดให้เป็น 2

อนุปริญญา/ปวส. กำหนดให้เป็น 3

ปริญญาตรี กำหนดให้เป็น 4

ปริญญาโท และสูงกว่า กำหนดให้เป็น 5

EM เป็นตัวแปรเชิงกลุ่ม คือ ระดับการศึกษาของมารดา

ประถมศึกษาหรือต่ำกว่า กำหนดให้เป็น 0

มัธยมศึกษาตอนต้น กำหนดให้เป็น 1

มัธยมปลาย/ปวช. กำหนดให้เป็น 2

อนุปริญญา/ปวส. กำหนดให้เป็น 3

ปริญญาตรี กำหนดให้เป็น 4

ปริญญาโท และสูงกว่า กำหนดให้เป็น 5

OF เป็นตัวแปรเชิงกลุ่ม คือ อาชีพของบิดา

ข้าราชการ กำหนดให้เป็น 0

พนักงานรัฐวิสาหกิจ และบริษัทเอกชน กำหนดให้เป็น 1

รับจ้าง กำหนดให้เป็น 2

ธุรกิจส่วนตัว กำหนดให้เป็น 3

พ่อบ้าน กำหนดให้เป็น 4

เกษตรกร กำหนดให้เป็น 5

OM เป็นตัวแปรเชิงกลุ่ม คือ อาชีพของมารดา

ข้าราชการ กำหนดให้เป็น 0

พนักงานรัฐวิสาหกิจ และบริษัทเอกชน กำหนดให้เป็น 1

รับจ้าง กำหนดให้เป็น 2

ธุรกิจส่วนตัว กำหนดให้เป็น 3

พ่อบ้าน กำหนดให้เป็น 4

เกษตรกร กำหนดให้เป็น 5

R เป็นตัวแปรเชิงปริมาณ คือ จำนวนเงินที่ได้รับต่อเดือน (บาท/เดือน)

PT เป็นตัวแปรหุ่น

ไม่มีรายได้เสริมจากการทำงานพิเศษ กำหนดให้เป็น 1

มีรายได้เสริมจากการทำงานพิเศษ กำหนดให้เป็น 0

SF เป็นตัวแปรเชิงปริมาณ คือ ค่าใช้จ่ายสำหรับอาหาร และเครื่องดื่ม (บาท/วัน)

L เป็นตัวแปรหุ่น

อาศัยอยู่บ้าน กำหนดให้เป็น 1

อาศัยอยู่หอพัก กำหนดให้เป็น 0

CM เป็นตัวแปรหุ่น

อาหารเช้า ปรุงอาหารรับประทานเอง กำหนดให้เป็น 1

ซื้ออาหารปรุงสำเร็จมารับประทาน/รับประทานอาหารนอกบ้าน กำหนดให้เป็น 0

CL เป็นตัวแปรหุ่น

อาหารกลางวัน ปรุงอาหารรับประทานเอง กำหนดให้เป็น 1

ซื้ออาหารปรุงสำเร็จมารับประทาน/รับประทานอาหารนอกบ้าน กำหนดให้เป็น 0

CE เป็นตัวแปรหุ่น

อาหารเย็น ปรุงอาหารรับประทานเอง กำหนดให้เป็น 1

ซื้ออาหารปรุงสำเร็จมารับประทาน/รับประทานอาหารนอกบ้าน กำหนดให้เป็น 0

KM เป็นตัวแปรเชิงปริมาณ คือ ปริมาณแคลอรี่ในอาหารเช้า (kcal/มื้อ/วัน)

KL เป็นตัวแปรเชิงปริมาณ คือ ปริมาณแคลอรี่ในอาหารกลางวัน (kcal/มื้อ/วัน)

KE เป็นตัวแปรเชิงปริมาณ คือ ปริมาณแคลอรี่ในอาหารเย็น (kcal/มื้อ/วัน)

F เป็นตัวแปรเชิงปริมาณ คือ ความถี่ในการบริโภคอาหารฟาสฟู้ด (ครั้ง/สัปดาห์)

DR เป็นตัวแปรหุ่น

ไม่บริโภคเครื่องดื่มประเภทน้ำอัดลม น้ำหวาน ชา กาแฟ กำหนดให้เป็น 1

บริโภคเครื่องดื่มประเภทน้ำอัดลม น้ำหวาน ชา กาแฟ กำหนดให้เป็น 0

E เป็นตัวแปรหุ่น

ไม่ออกกำลังกาย กำหนดให้เป็น 1

ออกกำลังกาย กำหนดให้เป็น 0

OP เป็นตัวแปรหุ่น

ไม่มีความเต็มใจจ่ายในการรักษาโรคอ้วน กำหนดให้เป็น 1

มีความเต็มใจจ่ายในการรักษาโรคอ้วน กำหนดให้เป็น 0

I เป็นตัวแปรเชิงปริมาณ

Incomplete Information and Asymmetric Information คำนวณจากแบบสอบถามส่วนที่ 2 คำถามเกี่ยวกับความรู้ความเข้าใจเกี่ยวกับโรคอ้วนที่มีผลกระทบต่อสุขภาพ ประกอบด้วยคำถามจำนวน 13 ข้อ คำถามแต่ละข้อ มีคะแนนเท่ากัน เท่ากับ 1 คะแนน นำคะแนนที่ได้หารด้วย 13

M เป็นตัวแปรเชิงปริมาณ

พฤติกรรมผู้บริโภคแบบ Myopic View หรือสายตาสั้น คำนวณจากแบบสอบถามในส่วนที่ 3 คำถามเกี่ยวกับการตระหนัก และให้ความสำคัญกับการมีสุขภาพที่ดีในอนาคต ประกอบด้วยคำถามจำนวน 11 ข้อ คำถามแต่ละข้อมีคะแนนเต็มที่ข้อละ 5 คะแนนเท่ากัน ดังนั้นคะแนนเต็มของคำถามในส่วนที่ 3 เท่ากับ 55 คะแนน นำคะแนนที่ได้หารด้วย 55

สามารถเขียนในรูปสมการ Logit Model เพื่อประมาณค่าพารามิเตอร์ (β) ด้วยวิธี Maximum Likelihood ดังนี้

$$\log\left(\frac{P(y=1)}{1-P(y=1)}\right) = \beta_0 + \beta_1 S + \beta_2 W - \beta_3 HI + \beta_4 FI - \beta_5 EF - \beta_6 EM + \beta_7 OF + \beta_8 OM + \beta_9 R + \beta_{10} PT + \beta_{11} SF + \beta_{12} L + \beta_{13} CM + \beta_{14} CL + \beta_{15} CE + \beta_{16} KM + \beta_{17} KL + \beta_{18} KE + \beta_{19} F + \beta_{20} D - \beta_{21} E - \beta_{22} OP - \beta_{23} I - \beta_{24} M$$

สมมติฐานของสมการ Logit Model มีดังนี้

1) เพศ การรับรู้น้ำหนักของตัวเอง การประกันสุขภาพเอกชน อาชีพของบิดา อาชีพของมารดา สถานที่พักอาศัย มีความสัมพันธ์ในทิศทางเดียวกันกับการเป็นโรคอ้วน การประมาณการค่าสัมประสิทธิ์มีเครื่องหมายเป็นบวก

2) รายได้ต่อเดือนของครอบครัว จำนวนเงินที่ได้รับต่อเดือน รายได้เสริมจากการทำงานพิเศษ ค่าใช้จ่ายสำหรับอาหาร และเครื่องดื่ม ลักษณะของอาหารที่รับประทาน (เช้า กลางวัน เย็น ปรุงอาหารเอง หรือรับประทานอาหารนอกบ้าน) ปริมาณแคลอรี่ในอาหารที่รับประทาน (เช้า กลางวัน เย็น) การบริโภคอาหารฟาสฟู้ด การบริโภคเครื่องดื่ม มีความสัมพันธ์ในทิศทางเดียวกันกับการเป็นโรคอ้วน การประมาณการค่าสัมประสิทธิ์มีเครื่องหมายเป็นบวก

3) การศึกษาของบิดา การศึกษาของมารดา การออกกำลังกาย ความเต็มใจจ่ายในการรักษาโรคอ้วน มีความสัมพันธ์ในทิศทางตรงข้ามกับการเป็นโรคอ้วน การประมาณการค่าสัมประสิทธิ์มีเครื่องหมายเป็นลบ

4) ข้อมูลข่าวสาร พฤติกรรมผู้บริโภคแบบ Myopic View มีความสัมพันธ์ในทิศทางตรงข้ามกับการเป็นโรคอ้วน การประมาณการค่าสัมประสิทธิ์มีเครื่องหมายเป็นลบ

3.2 ประชากร และวิธีการสุ่มตัวอย่าง

ประชากรในการศึกษาครั้งนี้ ได้แก่ นักศึกษาของมหาวิทยาลัยทั้งหมดที่อยู่ในจังหวัดเชียงใหม่ จำนวน 7 มหาวิทยาลัย ซึ่งประกอบด้วย 1) มหาวิทยาลัยเชียงใหม่ 2) มหาวิทยาลัยราชภัฏเชียงใหม่ 3) มหาวิทยาลัยราชมงคลล้านนา วิทยาเขตภาคพายัพ จังหวัดเชียงใหม่ 4) มหาวิทยาลัยฟาร์อีสเทอร์น 5) มหาวิทยาลัยแม่โจ้ 6) มหาวิทยาลัยนอร์ท – เชียงใหม่ และ 7) มหาวิทยาลัยพายัพ

กลุ่มตัวอย่างที่เป็นตัวแทนของการศึกษา คือ นักศึกษามหาวิทยาลัยทั้ง 7 แห่งในจังหวัดเชียงใหม่ ในปีการศึกษา 1/2552 ซึ่งเป็นทั้งนักศึกษาที่มีภาวะโรคอ้วน และไม่มีภาวะโรคอ้วน เป็นนักศึกษาทั้งเพศหญิง และเพศชาย ที่กำลังศึกษาอยู่ในระดับปริญญาตรี ชั้นปีที่ 1 – ชั้นปีที่ 4 ภาคปกติ อายุระหว่าง 18 – 21 ปี ซึ่งทั้ง 7 มหาวิทยาลัย มีนักศึกษาที่ลงทะเบียนเรียนทั้งสิ้นในปี พ.ศ. 2552 จำนวน 79,503 คน

โดยใช้ค่าดัชนีมวลกาย หรือค่า BMI เป็นเกณฑ์แบ่งกลุ่มตัวอย่างออกเป็น 2 กลุ่ม ได้แก่ กลุ่มที่ 1 นักศึกษากลุ่มผอม กับกลุ่มที่มีน้ำหนักปกติ ซึ่งมีค่า BMI น้อยกว่า 23 กิโลกรัมต่อตารางเมตร กลุ่มที่ 2 ได้แก่ นักศึกษาที่มีน้ำหนักเกิน นักศึกษาที่อ้วน และนักศึกษาที่อ้วนรุนแรง ซึ่งมีค่า BMI มากกว่าหรือเท่ากับ 23 กิโลกรัมต่อตารางเมตร โดยที่เกณฑ์พิจารณาค่า BMI กำหนดที่ 23

เนื่องจากการศึกษาคนในเอเชีย พบว่าสุขภาพจะดีที่สุด ถ้า BMI อยู่ระหว่าง 18.5 – 22.9 ซึ่ง BMI $\geq$ 23 มีความเสี่ยงต่อการเกิดโรคเรื้อรังต่าง ๆ และในประเทศไทยก็พบว่าผู้ป่วยโรคเบาหวาน ซึ่งมีความสัมพันธ์กับค่า BMI พบตั้งแต่ในคนที่มี BMI $\geq$ 23

การกำหนดขนาดของกลุ่มตัวอย่าง และการสุ่มตัวอย่าง แบ่งเป็น 3 ขั้นตอน ดังนี้

ขั้นที่ 1 คำนวณหาขนาดของกลุ่มตัวอย่างที่เหมาะสม

สูตรที่ใช้คำนวณหาขนาดของประชากร ได้แก่

แบบที่ 1 การใช้เกณฑ์ มีรายละเอียด ดังนี้

- ประชากรมีจำนวนในหลักร้อย ใช้จำนวนกลุ่มตัวอย่าง 15-30%
- ประชากรมีจำนวนในหลักพัน ใช้จำนวนกลุ่มตัวอย่าง 10-15%
- ประชากรมีจำนวนในหลักหมื่น ใช้จำนวนกลุ่มตัวอย่าง 5-10%

แบบที่ 2 ใช้ตาราง Krejcie & Morgan ซึ่งRobert V. Krejcie แห่งมหาวิทยาลัย Minisota และ Earyle W. Morgan แห่งมหาวิทยาลัย Texas (1970 : 608-609) ได้สร้างตารางขนาดประชากรและขนาดกลุ่มตัวอย่างไว้ ดังนี้

ตารางที่ 3.1 แสดงจำนวนประชากรและจำนวนกลุ่มตัวอย่างของ Krejcie and Morgan

ประชากร	กลุ่มตัวอย่าง	ประชากร	กลุ่มตัวอย่าง	ประชากร	กลุ่มตัวอย่าง
10	10	120	92	500	217
20	19	140	103	600	234
30	28	160	113	700	254
40	36	180	123	1000	278
50	44	200	132	1500	306
60	52	250	152	2000	322
70	59	300	169	5000	357
80	66	360	186	10000	370
90	73	400	196	50000	381
100	80	460	210	100000	384

ที่มา : Robert V. Krejcie and Earyle W. Morgan. Educational and Psychological Measurement, 1970 : 608-609

สำนักงานคณะกรรมการวิจัยแห่งชาติ ห้องสมุดงานวิจัย

แบบที่ 3 สูตรการคำนวณของ W.G.cochran ,1953 กรณีไม่ทราบขนาดประชากร

$$n = \frac{P\,(1-P)Z^2}{d^2}$$

เมื่อ n คือ จำนวนกลุ่มตัวอย่างที่ต้องการ

P คือ สัดส่วนของประชากรที่ผู้วิจัยต้องการสุ่ม (โดยทั่วไปนิยมใชสัดส่วน 30% = 0.30)

Z คือ ระดับความมั่นใจที่กำหนด หรือระดับนัยสำคัญทางสถิติ

- ที่ระดับความเชื่อมั่น 90% หรือ ที่ระดับนัยสำคัญ 0.10 , Z = 1.65
- ที่ระดับความเชื่อมั่น 95% หรือ ที่ระดับนัยสำคัญ 0.05 , Z = 1.96
- ที่ระดับความเชื่อมั่น 99% หรือ ที่ระดับนัยสำคัญ 0.01 , Z = 2.58

d คือ สัดส่วนความคลาดเคลื่อนที่ยอมให้เกิดขึ้นได้

- ที่ระดับความเชื่อมั่น 90% สัดส่วนความคลาดเคลื่อนเท่ากับ 0.10
- ที่ระดับความเชื่อมั่น 95% สัดส่วนความคลาดเคลื่อนเท่ากับ 0.05
- ที่ระดับความเชื่อมั่น 99% สัดส่วนความคลาดเคลื่อนเท่ากับ 0.01

แบบที่ 4 สูตรการคำนวณของ Poscoe, 1975 : 183 กรณีไม่ทราบขนาดประชากร

$$e = z\frac{s}{\sqrt{n}} \quad \text{และ} \quad n = \left(\frac{z*s}{e}\right)^2$$

เมื่อ e แทน ความคลาดเคลื่อนมากที่สุดที่ยอมรับได้

- ระดับความเชื่อมั่น 90% สัดส่วนความคลาดเคลื่อนเท่ากับ 0.10
- ระดับความเชื่อมั่น 95% สัดส่วนความคลาดเคลื่อนเท่ากับ 0.05
- ระดับความเชื่อมั่น 99% สัดส่วนความคลาดเคลื่อนเท่ากับ 0.01

Z แทน ความมั่นใจที่ระดับนัยสำคัญทางสถิติ

- ที่ระดับความเชื่อมั่น 90% หรือ ที่ระดับนัยสำคัญ 0.10 , Z = 1.65
- ที่ระดับความเชื่อมั่น 95% หรือ ที่ระดับนัยสำคัญ 0.05 , Z = 1.96
- ที่ระดับความเชื่อมั่น 99% หรือ ที่ระดับนัยสำคัญ 0.01 , Z = 2.58

S แทน ส่วนเบี่ยงเบนมาตรฐาน

n แทน ขนาดของกลุ่มตัวอย่าง

แบบที่ 5 สูตรการคำนวณของยามาเน่ (Taro Yamane,1973 : 125) กรณีที่ทราบขนาดประชากร

$$n = \frac{N}{1 + N(e)^2}$$

เมื่อ n แทน ขนาดกลุ่มตัวอย่าง

N แทน ขนาดประชากร

e แทน คลาดคลาดเคลื่อนของกลุ่มตัวอย่าง (ปกตินิยมระดับความเชื่อมั่น 95%)

- ระดับความเชื่อมั่น 90% สัดส่วนความคลาดเคลื่อนเท่ากับ 0.10
- ระดับความเชื่อมั่น 95% สัดส่วนความคลาดเคลื่อนเท่ากับ 0.05
- ระดับความเชื่อมั่น 99% สัดส่วนความคลาดเคลื่อนเท่ากับ 0.01

การศึกษาปัจจัยทางเศรษฐสังคมที่มีผลต่อการเกิดภาวะโรคอ้วนของนักศึกษาในมหาวิทยาลัย จังหวัดเชียงใหม่ใช้สูตรการคำนวณขนาดตัวอย่างของ Taro Yamane โดยมีระดับความเชื่อมั่นที่ 95% เนื่องจากทราบจำนวนประชากร และสูตรของยามาเน่ใช้กับข้อมูลที่เป็น Interval Scale และ Ratio Scale

Interval Scale คือ ข้อมูลที่มีลักษณะจำแนกกลุ่ม เรียงลำดับ และแบ่งเป็นช่วง ๆ ได้ เช่น คะแนนสอบ ซึ่งเป็นข้อมูลที่บอกความแตกต่างในเชิงปริมาณด้วยหน่วยที่เท่ากัน และระบุได้ว่าแต่ละคนมีความแตกต่างกันเป็นปริมาณเท่าใด ซึ่งแบบสอบถามในส่วนที่ 2 มีการวัดความรู้ความเข้าใจของนักศึกษา ซึ่งผลคะแนนจะระบุถึงความแตกต่างของความรู้ความเข้าใจเกี่ยวกับโรคอ้วน ที่ส่งผลให้นักศึกษาบางคนอ้วน บางคนผอมแตกต่างกัน

Ratio Scale คือ ข้อมูลที่จำแนกกลุ่ม เรียงลำดับแบ่งเป็นช่วง ๆ เท่า ๆ กัน สามารถเปรียบเทียบในเชิงอัตราส่วนได้ เช่น รายได้ ค่าใช้จ่าย น้ำหนัก ส่วนสูง ซึ่งเป็นข้อมูลในส่วนที่ 1 ของแบบสอบถาม

Ordinal Scale คือ ข้อมูลจำแนกกลุ่ม หรือประเภท และเรียงอันดับได้ เช่น ระดับความพึงพอใจ ระดับความคิดเห็น (มากที่สุด มาก ปานกลาง น้อย น้อยที่สุด) ซึ่งเป็นส่วนที่ 3 ของแบบสอบถาม

แทนค่า

$$n = \frac{79,503}{1+(79,503(0.05)^2)} = 397.99$$

ดังนั้น ตัวอย่างที่ใช้ในการวิจัยครั้งนี้ เท่ากับ 398 คน แต่ทำการเก็บข้อมูล 430 คน เพราะอาจมีแบบสอบถามบางฉบับที่ให้ข้อมูลไม่ครบถ้วน

ขั้นที่ 2 หาจำนวนประชากรที่ต้องการสุ่มในแต่ละมหาวิทยาลัย

การสุ่มตัวอย่างมี 2 แบบ ได้แก่

1) การสุ่มตัวอย่างแบบไม่อาศัยความน่าจะเป็น (Non Probability Sampling)

1.1) การสุ่มแบบบังเอิญ (Accidental Sampling)

1.2) การสุ่มแบบกำหนดโควต้า (Quota Sampling) เป็นการกำหนดกลุ่มย่อยตามต้องการ โดยอาศัยสัดส่วนขององค์ประกอบกลุ่มประชากรตามเพศ หรือการศึกษา

1.3) การสุ่มแบบเจาะจง (Purposive Sampling) เป็นการกำหนดตัวอย่างตามลักษณะที่ต้องการ

1.4) การสุ่มตามความสะดวก (Convenience Sampling)

1.5) การสุ่มตัวอย่างแบบก้อนหิมะ (snowball sampling) การโยนหินถามทาง

2) การสุ่มตัวอย่างแบบอาศัยความน่าจะเป็น (Probability Sampling) แบ่งเป็น 2 แบบ

2.1) ประชากรมีลักษณะคล้ายคลึงกัน (Homogeneous) ได้แก่

- การสุ่มตัวอย่างแบบง่าย (Simple Random Sampling) เป็นการให้ทุกหน่วยประชากรมีโอกาสถูกเลือกเท่า ๆ กัน เช่น การจับฉลาก การใช้ตารางเลขสุ่ม

- การสุ่มตัวอย่างแบบมีระบบ (Systematic Random Sampling) ใช้ในกรณีทราบหรือไม่ทราบขนาดประชากร โดยใช้การนับระยะห่างเท่า ๆ กัน แล้วสุ่มเลือกตัวอย่าง

2.2) ประชากรมีลักษณะต่างกัน (Heterogeneous) ได้แก่

- การสุ่มตัวอย่างแบบตามชั้นภูมิ (Statified Random Sampling) ใช้กับประชากรที่มีหลายลักษณะรวมกัน การสุ่มแบบนี้ทำให้ภายในกลุ่มมีความคล้ายคลึงกัน แต่ระหว่างกลุ่มมีความแตกต่างกัน

- การสุ่มตัวอย่างแบบกลุ่ม (Cluster Sampling) เป็นการแบ่งประชากรหลายลักษณะเป็นกลุ่ม ๆ แล้วสุ่มประชากรกลุ่มย่อยมาศึกษา 1- 2 กลุ่ม

- การสุ่มตัวอย่างแบบหลายขั้นตอน (Multi – Stage Sampling) ใช้ในกรณีที่มีประชากรจำนวนมาก ไม่สามารถทำรายชื่อทุกคนได้ ทำได้เฉพาะกลุ่มที่เลือกได้

การศึกษาปัจจัยทางเศรษฐสังคมที่มีผลต่อการเกิดภาวะโรคอ้วนของนักศึกษาในมหาวิทยาลัย จังหวัดเชียงใหม่ใช้การสุ่มตัวอย่างแบบชั้นภูมิตามสัดส่วน (Proportional Stratified Random Sampling) เนื่องจากประชากรมีลักษณะต่างกันในแง่พฤติกรรมการบริโภคอาหาร และกิจกรรมทางกาย ซึ่งเกิดจากการมีข้อมูลข่าวสารที่แตกต่างกัน และตระหนักถึงผลกระทบของโรค

อ้วนที่อาจเกิดขึ้นในอนาคตแตกต่างกัน แต่ก็มีความคล้ายคลึงกันในช่วงอายุ 18 – 21 ปี และกำลังศึกษาในระดับอุดมศึกษา

จำนวนตัวอย่างในแต่ละกลุ่ม/ชั้น = $\frac{\text{จำนวนตัวอย่างทั้งหมด} * \text{จำนวนประชากรในแต่ละชั้นภูมิ}}{\text{จำนวนประชากรทั้งหมด}}$

ตารางที่ 3.2 จำนวนประชากรที่เป็นกลุ่มตัวอย่าง แบ่งตามมหาวิทยาลัยในจังหวัดเชียงใหม่

ชื่อมหาวิทยาลัย	ประชากร	ตัวอย่าง
1. มหาวิทยาลัยเชียงใหม่	26,263	131
2. มหาวิทยาลัยราชภัฏ เชียงใหม่	23,706	118
3. มหาวิทยาลัยแม่โจ้	12,436	62
4. มหาวิทยาลัยราชมงคลล้านนา วิทยาเขตภาคพายัพ จังหวัดเชียงใหม่	7,567	38
5. มหาวิทยาลัยพายัพ	6,029	30
6. มหาวิทยาลัยฟาร์อีสเทอร์น	2,095	11
7. มหาวิทยาลัยนอร์ท – เชียงใหม่	1,407	8
รวม	**79,503**	**398**

ที่มา: สำนักงานคณะกรรมการการอุดมศึกษา กระทรวงศึกษาธิการ (2552)
http://www.ingo.mua.go.th/CE-app2552/INFO_UNIV/index.php

ขั้นที่ 3 สุ่มตัวอย่างแบบไม่อาศัยความน่าจะเป็น (Non Probability Sampling)

เมื่อผู้วิจัยแบ่งกลุ่มตัวอย่างตามมหาวิทยาลัยแล้ว คิดสัดส่วนต่อจำนวนประชากรของแต่ละมหาวิทยาลัยได้จำนวนกลุ่มตัวอย่างที่ต้องการแล้ว ก็ทำการสุ่มแบบบังเอิญ (Accidental Sampling) โดยให้กลุ่มตัวอย่างกระจายในเพศ ชั้นปี และคณะของนักศึกษาทั้ง 7 มหาวิทยาลัย เนื่องจากในการศึกษาครั้งนี้มีสมมติฐานว่า คณะ และชั้นปีการศึกษาของนักศึกษาไม่มีความแตกต่างกันอย่างมีนัยสำคัญต่อการเกิดภาวะโรคอ้วน

3.3 เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยครั้งนี้ ได้แก่ แบบสอบถาม โดยแบ่งเป็น 2 ประเภท คือ 1) แบบสอบถามเชิงปริมาณ และแบบสอบถามเชิงคุณภาพ

1) แบบสอบถามเชิงปริมาณ แบ่งออกเป็น 4 ส่วน ดังนี้

ส่วนที่ 1 เป็นคำถามเกี่ยวกับปัจจัยด้านเศรษฐสังคม พฤติกรรมสุขภาพ การรับรู้น้ำหนักของตัวเอง ความเต็มใจจ่ายในการรักษาโรคอ้วน เช่น เพศ ที่พักอาศัย นิยมการบริโภคอาหารที่ปรุงเอง หรือรับประทานอาหารนอกบ้าน ประเภทอาหารที่รับประทานอาหารท้องถิ่น หรืออาหารฟาสฟูดต์ และการออกกำลังกาย ลักษณะแบบสอบถามเป็นแบบให้เลือกตอบ และเติมตัวเลข

ส่วนที่ 2 เป็นคำถามเกี่ยวกับความรู้ความเข้าใจเกี่ยวกับโรคอ้วนที่มีผลกระทบต่อสุขภาพ สาเหตุของโรคอ้วน การมีกิจกรรมทางกายที่เพียงพอ และการบริโภคอาหารในสัดส่วนที่เหมาะสม มีความเข้าใจในการอ่านฉลากโภชนาการอย่างถูกต้อง เป็นแบบสอบถามชนิด ใช่ – ไม่ใช่ โดยมีหลักเกณฑ์การให้คะแนน ดังนี้ ตอบถูกให้คะแนน = 1 ตอบผิดให้คะแนน = 0

เมื่อทำการรวบรวมคะแนนที่ได้ทั้งหมดแล้วเอา 13 หาร จะได้ค่าความรู้ความเข้าใจเกี่ยวกับโรคอ้วนที่มีผลกระทบต่อสุขภาพที่มีค่าระหว่าง 0 ถึง 1

ส่วนที่ 3 เป็นคำถามเกี่ยวกับการตระหนัก และให้ความสำคัญกับการมีสุขภาพที่ดีในอนาคต ลักษณะคำถามเป็นแบบ Rating Scale โดยให้คะแนนเป็น 5 ระดับ คือ มากที่สุด มาก ปานกลาง น้อย น้อยที่สุด โดยกำหนดเกณฑ์การให้คะแนนในแต่ละระดับ ดังนี้ การตระหนักถึงการมีสุขภาพดีในอนาคต มากที่สุด = 5 คะแนน มาก = 4 คะแนน ปานกลาง = 3 คะแนน น้อย = 2 คะแนน และน้อยที่สุด = 1 คะแนน

เมื่อทำการรวบรวมคะแนนที่ได้ทั้งหมดแล้วเอา 55 หาร จะได้คะแนนของการตระหนักและให้ความสำคัญกับการมีสุขภาพที่ดีในอนาคตที่มีค่าระหว่าง 0 ถึง 1

ส่วนที่ 4 เป็นข้อเสนอแนะเกี่ยวกับปัญหาโรคอ้วน แนวทางการแก้ปัญหา ลักษณะของคำถามเป็นแบบเลือกตอบ และแสดงความคิดเห็นแบบปลายเปิด

2) แบบสอบถามเชิงคุณภาพ ประกอบด้วยคำถาม 2 ส่วน มีวัตถุประสงค์ ดังนี้ 1) เพื่อหาแนวนโยบายที่เหมาะสมกับการแก้ปัญหาภาวะโรคอ้วนของวัยรุ่น ในจังหวัดเชียงใหม่ 2) เพื่อนำเสนอรูปแบบของฉลากอาหารที่ระบุข้อมูลทางโภชนาการในลักษณะที่ตรงความต้องการของผู้บริโภค เพื่อทำให้ผู้บริโภคเกิดความใส่ใจ และตระหนักถึงความสำคัญของการอ่านฉลากก่อนการตัดสินใจเลือกซื้อผลิตภัณฑ์

แบบสอบถามเชิงคุณภาพใช้แบบสอบถามแบบมีโครงสร้าง สอบถามในผู้ที่มีส่วนได้ส่วนเสียกับโรคอ้วน ได้แก่ นักศึกษาที่เป็นโรคอ้วนในมหาวิทยาลัยทั้ง 7 แห่งในจังหวัดเชียงใหม่ จำนวน 30 คน ผู้มีส่วนเกี่ยวข้องกับการดำเนินนโยบายการแก้ปัญหาด้านสุขภาพของประชาชน ได้แก่ เจ้าหน้าที่กองทันตแพทย์ กระทรวงสาธารณสุข และเจ้าหน้าที่สาธารณสุขจังหวัดเชียงใหม่ รวม 3 ท่าน

3.4 วิธีการเก็บรวบรวมข้อมูล

3.4.1 ศึกษาจากแนวคิด ทฤษฎี เอกสาร และข้อมูลงานวิจัยที่เกี่ยวข้อง เพื่อเป็นกรอบแนวคิดในการวิจัย แล้วสร้างเครื่องมือการวิจัยที่เป็นแบบสอบถาม

3.4.2 ก่อนการนำแบบสอบถามที่สร้างขึ้นสำหรับการวิจัยไปใช้จริง ผู้วิจัยได้นำแบบสอบถามที่สร้างขึ้นไปทดสอบกับกลุ่มตัวอย่างจำนวน 30 คน ในมหาวิทยาลัยเชียงใหม่ โดยให้มีการกระจายตัวของกลุ่มตัวอย่างในเพศ คณะ ชั้นปี ที่นักศึกษากำลังศึกษาอยู่ และนำข้อมูลกลับมาวิเคราะห์คุณภาพของแบบสอบถาม ดังนี้

1) การหาความเที่ยงตรง (Validity) คือ การที่แบบสอบถามสามารถวัดได้ตามวัตถุประสงค์ มี 2 องค์ประกอบ ได้แก่

1.1) ความเที่ยงตรงเชิงเนื้อหา (Content Validity) โดยการนำเสนอแบบสอบถามต่อคณะกรรมการที่ปรึกษาวิทยานิพนธ์ของมหาวิทยาลัยธุรกิจบัณฑิต จำนวน 5 ท่าน ผู้เชี่ยวชาญของโครงการประเมินเทคโนโลยี และนโยบายด้านสุขภาพ (Health Intervention Technology Assessment Program: HITAP) จำนวน 3 ท่าน ประกอบด้วยผู้ที่ทำวิจัยด้านเศรษฐศาสตร์สุขภาพ 2 ท่าน และนักสถิติ 1 ท่าน และผู้ทำงานด้านการกำหนดนโยบายการแก้ปัญหาโรคอ้วนของกองทันตสาธารณสุข กรมอนามัย กระทรวงสาธารณสุข 1 ท่าน รวมทั้งสิ้น 9 ท่าน เพื่อประเมินความสอดคล้องของเนื้อหา และคำถามในแบบสอบถามว่าถูกต้องตรงตามวัตถุประสงค์ของงานวิจัย

1.2) ความเที่ยงตรงตามเกณฑ์ (Criterion) คำถามแต่ละข้อต้องวัดได้ตามมาตรฐานที่ต้องการ เป็นการใช้ค่าสถิติเพื่อหาความสัมพันธ์ของตัวแปรในส่วนที่เป็นคำถามแบบถูก – ผิด ได้แก่ สถิติ Pearson, Spearman, T – Test สำหรับการวิจัยครั้งนี้ใช้ Pearson Correlation ในการทดสอบความสัมพันธ์ของตัวแปร ซึ่งเป็นคำถามในแต่ละข้อว่ามีความสัมพันธ์กันหรือไม่ เพราะถ้าหากคำถามมีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติ ในทางปฏิบัติผู้ที่ตอบข้อนี้ถูก ก็จะตอบอีกข้อหนึ่งถูกเช่นกัน ดังนั้นจึงควรพิจารณาตัดคำถามข้อใดข้อหนึ่งออกไป

2) การหาความเชื่อมั่น (Reliability) คือ การที่แบบสอบถามซึ่งเป็นเครื่องมือในการวิจัยมีความคงที่ ไม่ว่าวัดกี่ครั้งก็ให้ผลลัพธ์เหมือนเดิม หรือใกล้เคียงกับของเดิม เป็นวิธีการหาความสอดคล้องภายใน Internal Consistency โดยค่าความเชื่อมั่นที่ได้แบ่งเป็นค่าความเชื่อมั่นของทั้งแบบสอบถาม และค่าความเชื่อมั่นรายข้อ มี 2 วิธี ได้แก่

2.1) Kuder – Richardson' Method ได้แก่ สูตร KR20 หรือ KR21 ใช้สำหรับแบบสอบถามที่เป็นคำถามแบบถูก – ผิด สูตรการคำนวณจะวิเคราะห์คุณภาพแบบสอบถาม 2 อย่าง คือ (1) ความเป็นปรนัย ได้แก่ คำตอบที่ระบุว่าถูก หรือผิด ไม่ว่าใครตรวจก็ได้คำตอบเช่นเดียวกัน

(2) ค่าความยาก – ง่ายของคำถาม ค่าความยากง่ายระดับปานกลาง คือ 0.2 – 0.8 และค่าความยากง่ายระดับดี คือ 0.4 – 0.6

นอกจากนี้คำถามแบบถูก – ผิดควรมีอำนาจการจำแนก (Discrimination) คือ คำถามสามารถแบ่งคนออกเป็น 2 กลุ่ม ได้แก่ กลุ่มเก่ง กับกลุ่มไม่เก่ง แบบสอบถามที่ดีต้องมีค่าอำนาจการจำแนก (r) $r > 0.2$

2.2) Conbach's Alpha Coefficient ใช้กับคำถามที่เป็นแบบ Rating Scale ซึ่งคำนวณหาค่าได้จากโปรแกรม SPSS ผลการคำนวณจะระบุอำนาจการจำแนกรายข้อ อำนาจการจำแนกที่ดีต้องมีค่าตั้งแต่ 0.2 ขึ้นไป ส่วนค่าความเชื่อมั่นที่ดีควรมีค่าตั้งแต่ 0.7 ขึ้นไป

นอกจากนี้ข้อคำถามแบบถูก – ผิด และคำถามแบบ Rating Scale ที่ดีควรเป็นคำถามที่ไม่มีการชี้นำให้ผู้ตอบเลือกตอบแต่คำตอบอย่างใดอย่างหนึ่ง เช่น ชี้นำให้ตอบแต่ "ถูก" หรือชี้นำให้ตอบแต่ "มากที่สุด" แต่คำถามทำให้เกิดการกระจายในทุกคำตอบ ซึ่งทดสอบการแจกแจงของข้อมูลโดยวิธี Kolmogorov – Sminov

3.4.3 นำแบบสอบถามที่สมบูรณ์แล้วไปสอบถามนักศึกษาที่ปีนกลุ่มตัวอย่าง จำนวน 398 คน ใน 7 มหาวิทยาลัยของจังหวัดเชียงใหม่

3.5 การวิเคราะห์ข้อมูล

3.5.1 แบบสอบถามเชิงปริมาณ ที่ได้จากการสอบถามนักศึกษา จำนวน 398 คน วิเคราะห์ประมวลผลข้อมูลในส่วนที่ 1 ปัจจัยด้านเศรษฐสังคม พฤติกรรมสุขภาพ การรับรู้น้ำหนักของตัวเอง ความเต็มใจจ่ายในการรักษาโรคอ้วน ส่วนที่ 2 ปัจจัยด้านการมีความรู้ความเข้าใจเกี่ยวกับผลกระทบของโรคอ้วนต่อสุขภาพ และส่วนที่ 3 การตระหนักถึงการมีสุขภาพที่ดีในอนาคต โดยใช้แบบจำลอง Logit Model ในโปรแกรม Eviews

3.5.2 แบบสอบถามเชิงคุณภาพ ได้แก่ แบบสัมภาณ์เจาะลึกแบบมีโครงสร้าง (Structured Interviewed Form) และแบบสอบถามเชิงปริมาณในส่วนที่ 4 ข้อเสนอแนะ นำข้อมูลที่ได้จากการสัมภาษณ์ และการสังเกตแบบมีส่วนร่วม (Participant Obeservation) ทำการวิเคราะห์เนื้อหา (Content Analysis) เพื่อนำไปสู่การนำเสนอนโยบายในระดับมหาวิทยาลัย และระดับจังหวัดตามความเหมาะสมเชิงพื้นที่ และความเป็นไปได้ในทางปฏิบัติ